

RICE UNIVERSITY

**Chess Performance under Time Pressure:  
Evidence for the Slow Processes in Speed Chess**

by

**Yu-Hsuan Chang**

A THESIS SUBMITTED  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE

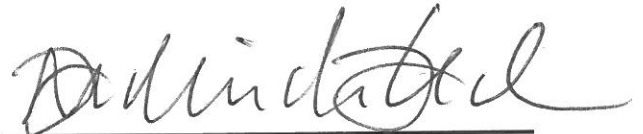
**Master of Arts**

APPROVED, THESIS COMMITTEE:



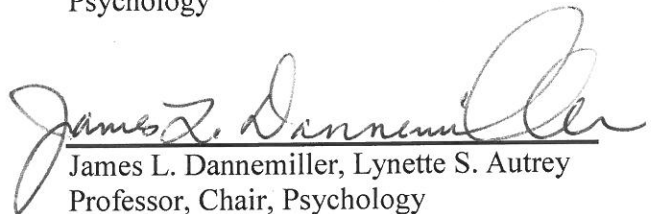
---

David M. Lane, Associate Professor,  
Committee Chair, Psychology, Statistics,  
and Management



---

Frederick L. Oswald, Professor,  
Psychology



---

James L. Dannemiller, Lynette S. Autrey  
Professor, Chair, Psychology

HOUSTON, TEXAS  
MAY 2013

## **ABSTRACT**

### **Chess Performance under Time Pressure: Evidence for the Slow Processes in Speed Chess**

by

**Yu-Hsuan Chang**

An influential theory of chess skill holds that expertise in chess is not due to greater depth of search by experts but, rather, to the ability to recognize familiar patterns of pieces. Although there is evidence that experts search deeper than non-experts, the data are not consistent. In this thesis, I propose “key-position theory” which states that only in a small number of key positions is it necessary to search deeply and it is these positions that experts search deeper than non-experts. Study 1 found, consistent with key-position theory, that the distribution of moves times is extremely skewed with some moves taking much longer than others. This pattern was more pronounced for the stronger players. Study 2 found that the errors made by weaker players involved less search than the errors made by stronger players. These findings suggest that search is an important component of chess expertise.

## **ACKNOWLEDGEMENTS**

I would like to express my sincere gratitude and thanks to my advisor Dr. David M. Lane for supporting me during these past two years. This thesis would not have been possible without his thoughtful guidance, constructive suggestions, and expert knowledge. I would also like to acknowledge my committee members, Dr. Fred L. Oswald and Dr. James L. Dannemiller, for their continued support and helpful suggestions on my thesis, and contributing their time generously in serving on my proposal and defense.

## TABLE OF CONTENTS

Acknowledgments.....	iii
Table of Contents.....	iv
List of Figures .....	v
List of Tables .....	vi
Introduction.....	1
Literature Review.....	2
Memory.....	2
Fast Processes in Move Choice.....	5
Summary of Fast Processes .....	10
Slow Processes in Move Choice.....	11
Summary of Slow Processes.....	20
Key-Position Theory.....	21
Reinterpreting Evidence for Recognition-Action Theory.....	22
Introduction to the Present Studies .....	25
Study 1: Move Times.....	27
Method .....	27
Time for Each Move .....	28
Results and Discussion .....	29
Study 2: Skill Differences in Poor Moves .....	37
Method .....	37
Poor Moves .....	37
Results and Discussion .....	41
General Discussion .....	44
References.....	47

## LIST OF FIGURES

Figure 1. <i>Time distribution of the moves from all subjects</i> .....	30
Figure 2. <i>The frequency distribution of the skewness of the time distribution from each subject</i> .....	30
Figure 3. <i>Time distribution of the moves from the stronger players</i> .....	32
Figure 4. <i>Time distribution of the moves from the weaker players</i> .....	32
Figure 5. <i>The boxplots of the two time distributions of the moves from the stronger players and the weaker players</i> .....	33
Figure 6. <i>The boxplots of the maximum of the time moves from the stronger players and The weaker players</i> .....	35
Figure 7. <i>The boxplots of the skewness of the time distributions from the stronger players and the weaker players</i> .....	35
Figure 8. <i>An example for Black's 1-ply poor move</i> .....	40
Figure 9. <i>An example for White's 2-ply poor move</i> .....	41
Figure 10. <i>The back-to-back dot plot of the n-ply poor move for the stronger and the weaker players</i> .....	43

**LIST OF TABLES**

Table 1. <i>The comparison of the time distributions of the moves for two different skill levels</i> .....	34
Table 2. <i>The proportion of n-ply poor moves over all moves for two different skill levels</i> .....	43

The study of chess expertise has attracted widespread attention since the seminal work of de Groot (1946). Over the past 70 years, researchers have proposed two distinct psychological mechanisms to explain how experts can make relatively good chess moves. These two distinct psychological mechanisms have been characterized by whether the information processing is fast or slow. Although most researchers agree that both fast and slow processes play at least some roles, there is disagreement about their relative importance.

*Fast Processes.* The vast majority of theory and research on the role of fast processes in chess has focused on a mechanism termed “recognition-action.” The recognition-action mechanism begins by quickly recognizing meaningful and familiar patterns in the chess positions so that preexisting patterns stored in long-term memory can be accessed. These “familiar perceptual structures that are evoked from long-term memory by the patterns on the board act as move generators” (Chase & Simon, 1973b, p. 268). This process is assumed to occur rapidly. Although there are other possible fast processes such as applying principles of chess (e.g., avoid double pawns) that may explain chess performance, they have received little or no research attention.

*Slow Processes.* Theory and research on slow processes have mainly focused on the process searching through the possible sequences of chess moves. These search processes have been referred to in various ways including “forward search” (Chabris & Hearst, 2003) and “look-ahead search” (Gobet & Simon, 1996b). There are other slow processes that are rarely discussed. One example is the possibility that there are some points in a game in which there is a choice of strategic directions which requires careful consideration. Samuel Reshevsky, a world-famous chess prodigy who eventually became

one of the world top players, stated that “[t]o a chess master, there is no such thing as an ‘obvious’ move.... Careful planning is the essence of chess strategy. Every move must be scrutinized with care. Each must be analyzed in light of plan under consideration” (1948).

## **Literature Review**

The structure of this literature review is as follows: I first review research on memory for chess positions. This literature is relevant to the problem at hand because recognition-action theory was derived from findings on chess memory. Next, I review research on recognition-action theory itself. Finally, I review literature addressing the relationship between expertise and depth of search.

### *Memory*

De Groot, considered by many as the father of chess psychology, designed a number of chess-related tasks which reflected specific cognitive functions and examined differences among players at different levels of skill on these tasks. The task showing the largest differences among players of different levels of expertise involved reconstructing briefly-presented board positions. There were four subjects in his study, including a grandmaster (Max Euwe, a world-champion), a weak master (de Groot himself), an expert player, and an average player. Across fourteen positions, both grandmaster (Euwe) and weak master (de Groot) were able to recall the board positions with 93% accuracy, whereas the two weaker players could recall only with 68% and 51% accuracy. De Groot and his colleagues (1964) carried out another crucial experiment in which subjects were required to recall and reproduce briefly presented boards with pieces arranged in meaningless ways. In other words, the pieces were arranged in a way that might have been impossible in actual games – there could be more than two knights, two bishops on



the same color square, etc. De Groot and his colleagues hypothesized that if the experts' superior memory is domain-specific, then the players' skill should not affect recall of randomized positions. The findings were consistent with their hypothesis. Therefore, de Groot (1965) stated that "a master is a master primarily by virtue of what he has been able to build up by experience; and this is: (a) a schooled and highly specific way of perceiving, and (b) a system of reproductively available methods, in memory" (p. 308).

De Groot and his colleagues considered the findings in random positions so obvious that they did not view it as worth publishing (Vicente & de Groot, 1990). Chase and Simon (1973a, 1973b) investigated de Groot's random position task and extend several of his other experimental tasks. They also found that the masters were able to recall meaningful chess positions much better than novices and replicated de Groot's finding that this expertise effect did not generalize to the randomly arranged chessboard.

Based on these findings, Chase and Simon (1973a, 1973b) developed the well-known "chunking theory" explanation of chess memory. According to this theory, chess players take pieces that are related in some meaningful patterns (e.g., color and proximity), and group them into "chunks" of four to five pieces. When a player views a board position, the player can recognize familiar chunks and label each chunk as a single unit. Chunks are well suited for working memory operations, and can be accessed quickly in long-term memory through their labels. Stronger players' superior memory for chess positions is hypothesized to be due to the size of the chunks rather than to the number of labels of chunks they can store in short-term memory.

Although the chunking theory explanation has been very influential, it has been challenged by some researchers as not fully adequate. According to chunking theory,

chess positions are perceived in terms of chunks, and those labeled chunks are held in short-term memory. However, Charness (1976) and Frey and Adesman (1976) found that skilled players' chess memory was unimpaired by various interpolated tasks. Therefore, the findings from these both studies indicate that the extracted information from a chess position presented for 5 s is immediately placed and held in long-term memory. In further support of this view, Frey and Adesman (1976) found that skilled chess players could recall two chess positions nearly as well as they recalled one. In 1979, Lane and Robertson applied the levels-of-processing theory of Craik and Lockhart (1972) to chess memory, and their results suggested that skilled chess memory is the result of a meaningful encoding in the long term memory. In the discussion, these authors suggested that recall is based on schemas (Rumelhart, 1977) of the position as a whole. Further supporting this view, Cooke, Atlas, Lane and Berger (1993) found evidence that high-level knowledge that is more abstract than perceived chunks play a role in chess memory, and that chunking theory has neglected the role of high-level knowledge. In their experiments, the skilled chess players were required to reconstruct the briefly-presented chess positions under two conditions: a high-level verbal description of the position was given either before or after its presentation. The authors hypothesized that if the high-level knowledge plays a role in the recalling of the chess position, then recall performance in the description-before condition should be better than the control description-after condition. The results were consistent with this hypothesis so that the authors concluded that "these results can be better explained in terms of high-level conceptual knowledge associated with each position than in terms of perceptual chunks" (p. 321).

In order to accommodate Cooke *et al.*'s (1993) results, Gobet and Simon (1996a) developed the "template theory" which expands on Chase and Simon's (1973a, 1973b) chunking theory. Gobet and Simon (1996a) proposed that "previously-stored multiple templates are used to remember chess positions.... A Grandmaster or Master holds in memory literally thousands of such patterns, each of which specifies the location of ten or a dozen pieces, with revisable defaults for others" (p. 31). Template theory differs from the explanation given by Cooke *et al.* in that it is based on pre-positioned pieces rather than schemas. In addition, each template consists of a core (the fixed information of recognized patterns) and slots (variable, nonessential information about other individual pieces). According to template theory, when multiple chess positions are presented to chess masters, template will be recognized mainly through the patterns of pieces from each board. The stronger players are better than the weaker players in recalling task because they have larger chunks and also those larger chunks are templates with slots.

#### *Fast Processes in Move Choice*

Based on the findings of skilled players' superior domain-specific memory, Chase and Simon (1973b) proposed what is often called the "recognition-action theory." According to recognition-action theory, expert chess players have very detailed information about chess positions stored in long-term memory, and this allows them to relate the position in the current game to previous games. The familiar perceptual structures in the board position allow the masters to "come up with good move almost instantaneously, seemingly by instinct and intuition" (Chase & Simon, 1973b, p. 269). In other words, the recognition-action mechanism allows recognition of patterns of pieces on the board, and these familiar patterns suggest possible moves by updating the internal

representation of board, which is called “mind’s eye.” Chase and Simon (1973b) explained the recognition-action mechanism by stating that “each familiar pattern serves as the condition part of a production. When this condition is satisfied by recognition of the pattern, the resulting action is to evoke a move associated with this pattern and to bring the move into short-term memory for consideration” (p. 269). That is, when a chess position is recognized by the players, it groups several chess pieces corresponding to chunks that are stored in long-term memory. These chunks are associated with plausible moves so that players can often immediately “see” the best move. According to recognition-action theory (Chase & Simon, 1973b; Simon & Gilmarin, 1973), the association of ability with brief recall is a central feature of different skilled performance. That is, the stronger players are able to access the larger amount of recognizable chess pattern than the weaker players. Therefore, Chase and Simon (1973b) concluded that “the most important processes underlying chess mastery are these immediate visual-perceptive processes rather than the subsequent logical-deductive thinking processes” (p. 215).

In order to model chess masters’ superior memory in multiple chess positions and data from expert problem solving, Gobet (1997) proposed SEARCH, a computational model based on CHREST. SEARCH model employs heuristics and pattern recognition to generate the possible moves by accessing a move or a sequence of move from a template or by selecting the next move from an heuristics. That is, masters narrow down and choose the best possible moves by rapidly accessing a template-based knowledge of chess. This “template-based knowledge” is based on a schematic structure rather than on the actual current board position. Due to the models that implement to chunking theory

and template theory, both theories stress the importance of recognition-action mechanisms.

According to recognition-action theory, the fast processes are responsible for expert move choice, and therefore the grandmaster's quality of move should not be much affected when the thinking time is limited. Testing this claim, Gobet and Simon (1996b) examined and compared the performance of Garry Kasparov, world-champion at the time, in regular tournament play and simultaneous chess. When playing simultaneous chess, Kasparov had much less time for his moves than did his opponents (four to eight relatively weaker players but still international masters). In the regular tournament, the players' first 40 moves have to finish in 2 hours; otherwise, they forfeit the game. The players in the regular tournament could spend considerable time on some moves, but have to move faster on other moves to make up the time. Therefore, Kasparov was allowed on average 3 minutes for each move (first 40 moves in 2 hours) in the regular tournament. In the simultaneous matches (e.g., six opponents), each of Kasparov's six opponents played the game under the regular tournament's rule (first 40 moves in 2 hours), while Kasparov has to play 240 moves within 2 hours. Thus, in the simultaneous chess, Kasparov played with a limit of 3 minutes (on average) for each round- that is, for each four to eight moves. Based on Kasparov's performance under different time constraints in regular tournament and simultaneous matches, the authors suggested Kasparov's rating was only slightly lower (100 Elo points) when he had to play faster in the simultaneous chess (from 2,750 to 2,646). Therefore, Gobet and Simon (1996b) stated that "recognition, by allowing knowledge to be accessed rapidly, allows the slower look-

ahead search to be greatly abridged or even dispensed with entirely without much loss in quality of play” (p. 53).

However, Kasparov prepared for these matches by studying the previous games of his opponents in detail to identify his opponents’ strength and weakness (1993a; 1993b; 1993c; 1993d; 1994). Based on his study of his opponents, he would have been able to prepare specific moves to steer the play into a specific positions that is not suit a particular opponent, so that he will be able to increase the likelihood of the opponents’ making mistakes. In other words, Kasparov’s performance in the simultaneous chess was not simply only replying on the fast processes.

Lassiter (2000) argued that the reason that Kasparov’s performance dropped by about 100 Elo points in simultaneous chess could be partly because of the reduction of opportunities and time he had to search for and evaluate future moves. Lassiter (2000) supported this by comparing the performance of expert players competing against chess-playing computers. Computers gained about 100 Elo points relative to their human opponent in 25-minute games (30 s per move, on average), and gained 200 or more Elo points in 5-minute games (6 s per move, on average). Lassiter’s (2000) explanation of this finding was that “the tendency for chess-playing computers to become relatively stronger at shorter time controls is most likely due to the fact that a human’s ability to engage in search-evaluation is more hampered by increasingly higher time constraints than is a computer’s” (p. 172). Gobet and Simon (2000) responded to Lassiter’s argument, claiming that his argument was unfounded, and that “the human difficulty could as readily be dearth of pattern-recognition time as dearth of search time, or time available to consult stored opening variations” (p. 174). However, this statement seems to contrary to

the previous statement that the recognition-action mechanism “provides a rapid index to the master’s chess knowledge and tolls of analysis” (Gobet & Simon, 1996b, p. 52). In addition, in this article, they acknowledged that search plays a role in the mechanisms of chess cognition, which ran contrary to the previous popular sentiment from the pattern recognition camp that forward search played a minimal role.

Consistent with Gobet and Simon’s (1996b) claim the stronger players’ quality of moves are only slightly affected by the reduction of the thinking time in the simultaneous chess, Calderwood, Klein, and Crandall (1988) found little difference in quality of moves between tournament chess and speed chess. However, they found that an interaction between skill and game type, such that the presence of time pressure decreased the quality of play of relatively weaker players, but not the quality of stronger players. The authors also concluded that rapid pattern-recognition process played a more important role in the skill of the stronger players than in the weaker players. They stated that “these results were interpreted as supporting the view that the more highly skilled players are able to rely more extensively on rapid recognitional processes than less skilled players” (p. 481).

Burns (2004) analyzed archival data on both regular and speed chess (5-minutes) tournament games. Burns (2004) assumed that due to the reduction of the opportunity to use slow processes in speed chess, speed chess differences are due almost entirely to differences in fast processes. It follows from this assumption that if speed chess performance is highly correlated with regular chess performance, then most the variance in regular chess performance must be due to differences in fast processes. Consistent with his hypothesis, the players’ speed-chess performances were highly correlated with their

chess skill in the tournament chess. Specifically, Burns (2004) concluded that up to 81% of the variance in chess skill was accounted for by the fast process of pattern recognition, which was used in both types of chess game. Burns (2004) concluded that “the results are consistent with the finding that search does not improve much as players’ skill levels rise, which implies that search should not account for much of the variance in players’ performance” (p. 446). These studies have supported the views from Gobet and his colleagues that “recognition, by allowing knowledge to be accessed rapidly, allows the slower look-ahead search to be greatly abridged or even dispensed with entirely without much loss in quality of play” (Gobet & Simon, 1996b, p. 53).

#### *Summary of Fast Processes*

Historically, proponents of recognition-action theory have discounted the importance of forward search in chess cognition. Both chunking theory and template theory suggest that the concept that players use preexisting information from long-term memory to quickly find good moves. The foundation of recognition-action mechanism is based on evidence that stronger chess players seem to play based more on perception and memory rather than on forward search. One argument that has been used to support this view is that many studies have shown that there is little or no difference in maximal depth of search in given chess positions between stronger players and weaker players, especially when the experts above a certain skill level (Charness, 1981; Charness, 1989, Charness, 1991; Connors, Burns, & Campitelli, 2011; de Groot, 1965; Gobet, 1998). A second argument is that speed chess performance correlates highly with tournament chess performance. Since there is presumably no time for slow processes in speed chess, this suggests that the fast processes that account for this high correlation (Burns, 2004). A



third argument is the evidence provided by Gobet and Simon (1996b). Since the top player's performance against six to eight weaker players when playing was only slightly worse when he played normal chess. This result supports the importance of fast processes. These arguments are based on the findings that there was not much difference in forward search in experts above a certain skill level, and that time pressure did not greatly affect the performance of expert players.

#### *Slow Processes in Move Choice*

Research and theory on slow processes have mainly focused on the process of searching through the possible sequences of chess moves. Some researchers claimed that expert chess players have very detailed information about chess knowledge stored in long-term memory, and this allows them to search and evaluate the possible moves more efficiently and correctly after years of practice. This notion was first tested and then developed by Holding and his colleagues. From 1979 to 1985, Holding and his colleagues provided evidence that the main determinants of chess skill include: 1) the efficiency of search, 2) the accuracy of evaluation, and 3) the breadth of knowledge. Holding's main argumentation is that strong players search deeper than weaker players. However, unlike the consistent findings of strong players' superior memory, the findings with regard to depth of search are less consistent. One parameter of depth of search is called "maximal depth of search" which represents the greatest depth could be reached during the search in half-moves (plies). One of de Groot's well-known methods is the thinking-aloud protocol, which is essentially a verbal protocol applied to chess. In this task, subjects were first presented with a chessboard with pieces arranged in various positions for a short period (2–15 seconds), which were then cleared. The subjects were

then required to verbalize their thought processes for their next moves. The chess skill of the subjects ranged from Grandmaster to Candidate Master and one chess position was used. De Groot (1965) did not find differences between stronger (including world-champion) and relatively weaker players in either breadth or the depth of their move calculations. He did find expertise differences in the speed of their analysis. This difference between stronger and relative weaker players in their speed of analysis has implied that fast processes may play an important role in speed chess. It also could partially explain Calderwood *et al.*'s (1988) finding that time pressure is more detrimental to weaker players than to stronger players. Based on the finding of players' memory and depth of search, de Groot (1946, 1978) concluded that perception and memory were the main differentiators between experts and grandmasters, rather than the ability to look ahead in search of good moves.

Many studies have replicated de Groot's study with the same board position "A," which was specific mentioned in detail in de Groot's (1965) book. For instance, Gobet (1998) found that there was no difference in the maximal depth of search among Masters ( $M = 9.1$ ,  $SD = 3.8$ ), Experts ( $M = 8.9$ ,  $SD = 3.6$ ), class A players ( $M = 9$ ,  $SD = 6.5$ ), and class B players ( $M = 6.1$ ,  $SD = 3.1$ ). However, an analysis not reported by Gobet that I performed is that the comparison of the Master's and class B player's depth of search from the data. A t-test of the Master's and class B player's depth of search based on the standard deviation and mean provided in the article shows that Masters searched significantly deeper than class B players,  $t(22) = 2.12$ ;  $p = .0456$ . Similarly, Connors *et al.* (2011) found suggestive evidence that the masters search deeper even though the difference was not significant. Bilalic, McLeod, and Gobet (2008) compared the

difference between grandmasters and novices. The grandmasters' data of maximal depth was from de Groot's (1965) study, whereas the data of novices was from Gruber's (1991) study. Since both studies applied the same board position, Bilalic *et al.* (2008) compared these two groups and concluded that grandmasters ( $M = 7.4$ ,  $SD = 2.6$ ) reached substantially greater maximal depths than novices ( $M = 3.3$ ,  $SD = 1.8$ ).

Some studies have investigated the depth of search with different positions. In 1981, Charness designed an experiment in which the subjects were given 10 seconds to evaluate one endgame, one quiet middle, one tactical middle, and one tactical opening game position. Charness analyzed 136 thinking- aloud protocols from players ranging from 1283 to 2004 Chess Federation of Canada (CFC) rating points. Stronger players exhibited greater search depths, with maximum depth increasing by nearly 1.5 plies for every standard deviation of skill rating. Therefore, Charness (1981) suggested that as players increased in skill, the depth of their search increased, but this function possibly reached an asymptote once players reached the Elo rating of 2100. In order to test this hypothesis, Charness (1989) retested one subject, DH, with the four identical positions to those given in 1978, who advanced (in power law fashion) from 1600 to 2300 rating points after nine years (1978-1987), and found that there was no significant increase in the maximal depth of search in overall four positions. However, DH showed a tendency to search less deeply for the ending and quiet middle positions (from 6 plies in 1978 to 5 plies in 1987), whereas he showed a tendency to search more deeply for the tactical positions (from 7 plies in 1978 to 9 plies in 1987). Based on this result, in 1991, Charness asserted that this asymptote might actually occur far earlier, at the Elo rating of 1600. Based on Charness' (1981, 1989, 1991) findings, he concluded that the ability of the

depth of search increased by nearly 1.5 plies for every standard deviation of skill rating, which reached an asymptote at the rating of 1600.

Saariluoma (1990) examined the mean depth of search for players ranging from 1900 to 2500 Elo. In this experiment, the subjects were presented with a single tactical position and were required to generate a move with a 10-minute limit. Six positions were randomly selected in this study. The average depth of search of International Masters (3.6 moves) and Class-B players (4.6 moves) was less lower than that of Masters (5.1 moves).

Campitelli and Gobet (2004) presented subjects of various skill levels three complex chess positions, each requiring considerable forward search and evaluation, with a 30-minute limit for each position. The solutions for these three positions required 52 nodes and search to a depth of 23 plies, at least 60 nodes and a depth of 25 plies, and at least 70 nodes and a depth of 35 plies. They found that there was difference in the maximal depth of search across these three positions among grandmaster ( $M = 25$ ,  $SD = 12.2$ ), international master ( $M = 23.7$ ,  $SD = 5.1$ ), candidate master ( $M = 17.7$ ,  $SD = 4.9$ ), and class B player ( $M = 10.5$ ,  $SD = 2.2$ ). This supports the argument that stronger players are able to think deeper, faster, and also generate more nodes than weaker players.

Bilalic, McLeod, and Gobet (2009) examined how specialization in different openings affects expert chess players' performance of breadth and depth of search by presenting two types of position to two types of players (those specializing in the French defense and those specializing in the Sicilian defense). There was an interaction between player type and position type for depth and breadth searches. Players were able to search deeper when the position was within a player's area of specialization than when it was outside. In addition, there were three different levels of skills for each type of players:

international and grandmasters, masters, and candidate masters. The international and grandmasters searched significantly greater depth than masters, but there was no significant difference between other skill levels.

In sum, these data are inconsistent regarding expertise differences in maximal depth of search. However, as the above review indicated, the long-standing claim of a lack of relationship between chess skill and depth of search is not generally true and appears to hold for some but not all positions.

There is a role for slow processes other than depth of search. In 1982, Holding and Reynolds asked their subjects of skill levels ranging from novice to expert to think of the best move from three pseudo-random positions. They used pseudo-random positions to eliminate the recognizable configurations which are associated with possible moves as suggested by recognition-action theory. Chess skill was positively correlated with the quality of move selection, but was not correlated with the number of pieces reconstructed. Therefore, the authors argued that this difference as a function of strength that cannot be explained by recognition-action theory, since there is nothing familiar to recognize. Holding and Reynolds (1982) claimed that their results are “strongly against...Chase and Simon’s (1973a) interpretation of the relationship between playing strength and recall” (p. 240). In addition, Holding and Reynolds (1982) argued that Chase and Simon’s (1973a) recognition-action model is premature because this model derives from the findings on recall scores and it postulates that the move was played based on the players’ memory that associated with plausible move with recognized sub-patterns. Therefore, Holding and Reynolds’ (1982) results are often cited as evidence that pattern recognition and forward search are dissociated.

In 1985, Holding and Pfau asked chess players with different skill levels to numerically evaluate positions from actual games as moves were dictated. The subjects were presented with a middle game position and asked to evaluate the board from one player's perspective on a numerical scale, and then dictated the next three pairs of moves without changing the positions on the board. After each pair of moves was dictated, the subjects were asked to provide a new evaluation from their imagination of the new chess position. Holding and Pfau (1985) compared evaluations where the subject could see the final state of the board at the end of the 3 pairs of moves with evaluations where the subject had to imagine the previous 3 pairs of moves. Throughout all skill levels, evaluations were more accurate for moves closer to the final position of the game. They found that stronger players ( $> \text{Elo } 1500$ ) were better at evaluating imagined positions than weaker players ( $< \text{Elo } 1500$ ), and the discrepancies between the evaluations of actual and imagined positions were larger for weaker players across the three pairs of moves. Holding (1979, 1985) also investigated the way players evaluated the nodes of their search trees and again demonstrated that strong players evaluated positions better than weaker players. Thus, Holding concluded that "thinking ahead, in all its complexity, defines skill at chess" (1985, p. 256), and "chess play is based on the anticipated consequences of moves rather than on recognizing specific patterns" (1992, p. 16).

According to recognition-action theory, a top chess player's skill should not deteriorate much when the thinking time is substantially reduced (Gobet & Simon, 1996b), because the fast processes are more automatic and effortless and, thus it does not rely on time. Chabris and Hearst (2003) argued that a finding of more errors and more serious errors in speed chess than in tournament chess would call into question Gobet and

Simon's (1996b) view about the relative importance of pattern-recognition and calculation processes in chess skill. To test this, Chabris and Hearst (2003) applied a computer program, Fritz 5 chess-playing program (ChessBase, Hamburg, Germany), to identify the major blunders and obtain objective scores in three different game conditions: classical, rapid, and rapid blindfold games. Data from the rapid and rapid blindfold games were taken from six editions of the Monaco tournament, where top-level grandmasters played at about the same speed (less than 30 s per move, on average) with seeing and unseeing the actual board positions. The researchers looked specifically at blunders in 1,188 games. Each move in these games was numerically evaluated in pawns, a common measure of move quality in chess, and compared with the program's choice for the best move. A move evaluated as at least 1.5 pawns worse than the program's choice was considered a candidate blunder, and the move was declared as a true blunder unless the player moving had a 3.0 pawn advantage or higher even after the move. The grandmasters averaged 5.02 true blunders per 1,000 moves in classical games, 6.85 blunders per 1,000 moves in rapid games, and 7.63 blunders per 1,000 moves in blindfold rapid games. The authors concluded that grandmasters make fewer and smaller blunders when they play more slowly. This finding led Chabris and Hearst (2003) to state that "when more time is available, skilled players have both a greater opportunity to recognize more patterns as well as to analyze ahead" (p. 644). However, Gobet (2003) criticized Chabris and Hearst (2003) in an unpublished critique for several fallacies. Gobet (2003) questioned Chabris and Hearst's (2003) calculation of a 36.5% higher rate of blunders in rapid games over classical (from 5.02 to 6.85 per 1,000 moves), proposing that the true increase is 0.183%, the difference between the two blunder rates. Chabris and Hearst

(2005) argued against this claim, suggesting that Gobet (2003) was trying to minimize the effect they observed with inappropriate reasoning. Gobet (2003) found the difference between the rates of blunders, but claimed that this simple difference was the difference of interest. However, Chabris and Hearst (2003) pointed out that the difference is small because the rate of blunders in general for grandmasters is so small. To Chabris and Hearst, the difference of interest is the increase of blunders in rapid games compared to the number of blunders in classical games, and grandmasters make 36.5% more blunders in rapid games than classical ones. Gobet (2003) also argued that he shared a similar message with Chabris and Hearst (2003): “reduction in thinking time leads to loss in equality of play,” but “the disagreement lies in the magnitude of this loss, and its consequences for theories of expertise” (p. 12). However, Chabris and Hearst (2005) responded that a loss in quality of play due to a reduction of thinking time supports forward search and opposes pattern recognition. They claimed that a large enough difference in blunders between rapid and classical chess should invalidate the recognition-action process, and “the degree of opportunity to ‘think ahead’ appears very important” (Chabris & Hearst, 2003, p. 644).

In 2007, van Harreveld, Wagenmakers, and van der Maas analyzed the data from an online chess server, Internet Chess Club (ICC), which includes many of the best players in the world, with a total of over 30,000 active members. On ICC, players can play chess under 5 different time conditions: bullet (1-2 min), blitz (3-10 min), standard (10-60 min), 5-minutes, and 1-minute games. Players receive an individual stable rating after they have played a sufficient number of games for each of these options. In this study, subjects were randomly selected from the four different title categories in ICC



(listed in decreasing order of skill): Grandmaster, International Master, Master, and untitled players. For each of the four title categories, the authors examined the correlations between 1-minute and 5-minutes ratings, blitz and bullet ratings, and blitz and standard ratings. According to the hypothesis of recognition-action theory, the main differentiator between stronger and weaker players depends on their ability to recognize meaningful patterns, which are evoked from memory previously learned. Based on its hypothesis, one would expect that correlations between ratings of various time controls should increase with playing strength. However, their results showed that all three correlations decreased with players' strengths, which is opposing with the hypothesis of recognition-action theory. Therefore, van Harreveld *et al.* (2007) concluded that these findings posed a challenge for the importance of recognition-action, and slow processes are important for both strong players and weak players.

In 2012, Moxley, Ericsson, Charness and Krampe asked chess players with two different skill levels to generate best moves from three different difficulty levels of the tactical chess positions. The subjects were presented with a position and asked to generate best move. They analyzed the first move considered and the move chosen after five minutes. Across all levels of position difficulty, the quality of the final selected move was better than the first mentioned move, and this improvement was larger for the stronger players in the harder positions. An important finding is that on these harder positions, there was no skill difference in move quality of the first-mentioned move but a sizeable difference in move quality on the final move. Accordingly, Moxley et al. (2012) concluded that “[f]or the difficult problems... [there is] a clear superiority for the experts’ selected move attained through deliberation” (p. 76).

### *Summary of Slow Processes*

In sum, the role of slow processes in chess-playing skill was noted by Holding (1985), and studied further by others (e.g., Chabris & Hearst, 2003; Campitelli & Gobet, 2004; van Harreveld *et al.*, 2007). The foundation of the forward-search explanation of skill differences is based on the thesis that stronger chess players are able to search deeper and faster than weaker players. A first argument that has been used to support this view is that some studies have shown that there is a difference in maximal depth of search in certain given chess positions, not all positions, between the stronger players and weaker players (Campitelli & Gobet, 2004; Bilalic *et al.*, 2009) or novice (Bilalic *et al.*, 2008). A second argument that has been used to support this view is the finding that the stronger players are able to evaluate a position better than weaker players (Holding & Plau, 1985). A third argument that has been used to support this view is that stronger players play better quality moves selection and fewer blunders than the weaker players either with no time pressure or under time pressure (Campitelli & Gobet, 2004; Chabris & Hearst, 2003; Holding & Reynolds, 1982; Moxley *et al.*, 2012). Thus, it can be concluded, based on the above findings and arguments, that when the position requires the player to search deeply in order to find the best move, stronger players are capable to search deeper than weaker players. Consistent with Kasparov's (2010) statement: "[how many moves ahead he sees] depends on the position and how much time I have" (p. 205). In addition, Kasparov (2010) point out that he even visualized the winning position a full fifteen moves ahead in the game at the 1999 Hoogovens tournament. It is extremely unlikely that many players would have been capable of looking that far ahead. Not only search deeper when the position requires, Kasparov also well prepared for his opponents.

As described previously, Kasparov's plays involve long-term planning, which definitely cannot be accounted for by fast processes.

### *Key-Position Theory*

In this thesis, I propose "key-position theory" which states that slow processes are required in only a few key positions. The slow processes include not only depth of search, but also breadth of search and planning. The depth of search refers how far ahead the player looked, whereas the breath of search refers how many possible moves along the tree of possible variations the players considered. In addition, there are some points in a game in which a strategic decision must be made. These decisions require careful consideration and may take a relatively longer time. These types of positions that involve slow processes are all considered key positions. Unfortunately, there are no clear criteria to determine objectively which positions are the key positions. Determining objective criteria would be desirable but to do so is very complicated and beyond the scope of this thesis. Since the depth of search has been considered as the primary slow processes, this thesis focuses on it.

Key-position theory posits that skill differences in depth of search occur primarily in these positions and that stronger players do more calculations and may spend more time in these positions than do the weaker players even if they calculate faster. In addition, key-position theory posits that the stronger players are able to search more deeply than the weaker players even in speed chess or simultaneous chess because they only do the time-consuming search in key positions. It may be true that the calculation time is reduced under time pressure, but this does not mean that it eliminates all slow processes, especially these occur only in some critical positions. For example,

Viswanathan Anand, a world champion in speed chess (5 minutes for each player), spent 1 minute and 43 seconds on 4<sup>th</sup> move, when he played against Grandmaster Ilya Smirin at the Intel World Speed Chess Tournament in 1994. This shows clearly that slow processes are involved even in speed chess. In addition, I argue below that key-position theory provides alternative explanations for the major findings which are used to support recognition-action theory.

*Reinterpreting Evidence for Recognition-Action Theory*

A key result supporting recognition-action theory is the finding of no difference in the depth of search between the stronger and the weaker players (de Groot, 1965; Gobet, 1998; Connors *et al.*, 2011). However, these findings are not consistent across studies. First, a more powerful test of the difference between Master's and class B player's depth of search than the one used in Gobet's (1998) results in a significant difference between these two groups with Masters searching deeper than class B players. Second, Campitelli and Gobet's (2004) study found that the stronger players were able to think deeper, faster, and also generate more nodes than the weaker players. Third, Bilalic *et al.* (2009) found that if a position was within a player's area of specialization, the player was able to search deeper than if the position was outside the player's area of specialization.

Key-position theory provides an explanation of these inconsistent findings. According to key-position theory, the differences between studies in differences in depth of search occurred because only small percentage of positions requires the depth of search. Therefore, the studies not finding a skill difference in search may have not included any of these key positions.

A second argument for recognition-action theory is the finding that strong players play well even under severe time constraints. Proponents of this view argue that time pressure essentially eliminates slow processes such as forward search and planning. As a result, the high-quality play often found under time pressure must be due almost exclusively to fast processes. There are three studies based on this argument claiming support for recognition-action theory. First, Gobet and Simon (1996b) found that when the world champion, Kasparov, had to play quickly in simultaneous chess, his rating was only slightly lower than normal. Second, Calderwood *et al.* (1988) found that stronger players generated no difference in the average move quality between the regular and speed chess; on the contrary, the weaker players' move quality decreased when they played speed chess compared to regular chess. Calderwood *et al.* (1988) argued that the stronger players rely more on the fast recognition-action processes than weaker players because they were able to generate the high quality moves when they have to play fast in the speed chess. Third, the assumption that forward search plays at most a trivial role in speed chess is the basis of the argument made by Burns (2004). Burns (2004) argued that the high correlation found between the performances under time pressure and under less severe time constraints indicates that expertise, even when players have sufficient time, is largely due to fast processes.

In the followings, I critique these findings and discuss the interpretation of these results in terms of key-position theory. Regarding the performance of Kasparov in Gobet and Simon's (1996b) study, I present three issues that complicate the interpretation of their findings. First, many studies have questioned whether Kasparov played almost as well as his Elo rating during the simultaneous games (see Chabris & Hearst, 2005;

Lassiter, 2000). Second, since the authors did not provide the time that Kasparov spent on each individual move, there is not enough information to determine whether Kasparov calculated on some key positions or not. Third, as described previously, Kasparov is well-prepared for his opponents by studying the previous games of his opponents. His good results may have stemmed, in part, from his identifying weaknesses in his opponents' opening play and understanding of various types of positions and steering the games toward these openings and types of positions. Therefore, it is hard to interpret Kasparov playing so well in the simultaneous games solely in terms of recognition-action theory.

Recall that Calderwood *et al.* (1988) found no evidence that stronger players' move quality was lower in speed chess than in regular chess. This anomalous finding calls their conclusions into question. Moreover, these authors did not consider that the poor performance of the weaker players in speed chess may have been due to their slow calculation speed which would greatly reduce their move quality when time is severely limited.

In Burn's (2004) study, the high correlation between the stronger players' performances under time pressure and under less severe time constraints can be explained by key-position theory as follows. Key-position theory states that calculation is important in a small number of key positions and that the ability to play the correct move in these positions can determine the outcome of the game. Even in speed chess, these players may take enough time in these positions to do considerable calculations. If the ability to calculate in these positions is correlated with the ability to calculate in standard chess, then the correlation between players' ratings in the two types of chess would be based, in

part, on calculation ability. This contradicts the assumption of Burns (2004) that there is not time in speed chess for calculations to play a meaningful role.

An important limitation of these studies purportedly supporting recognition-action theory is that none of them measured times of individual moves. Therefore, it is unknown whether the players took sufficient time on some moves to determine if the calculations took place or not.

In sum, according to key-position theory, the depth of search and skill differences in depth of search depends on the positions. This slow process is used even in speed chess when total time is severely limited as long as it is possible to spend a relatively long time on a few moves.

#### *Introduction to the Present Studies*

As described above, although there have been several studies of the effect of time pressure on chess playing, these studies have only considered the average time per move and have not considered the time taken on individual moves. However, a recent study by Sigman, Etchemendy, Slezak, and Cecchi (2010) considered the time taken on individual moves in speed chess. They studied games in which each player had a total of 3 minutes for all of his or her moves. They found that response times were much shorter in the opening and endgame, and much longer in the middle game. This is not surprising since, as the authors noted, it is likely that the opening moves were played from memory and the end-game moves were played when there was little time left on the clock. This pattern was more pronounced for stronger than for weaker players, perhaps indicating that players budget their time differently depending on their playing strength. It is also possible that stronger players have memorized far more opening moves and techniques

for end-game play and thus could move quickly in those phases of the game. The data also showed that the move-time distributions were positively skewed. However, they interpreted this as being consistent with the general finding that response times are positively skewed. They did not consider the implications of the distribution shape on theories of chess skill. Moreover, they did not focus on differences among skill levels except in terms of differences among phases of the game. One interesting finding reported by Sigman *et al.* (2010) is that “higher rated players deploy time in a more variable manner, probably depending on necessities of the position” (p. 5). In terms of key-position theory, this implies that stronger players take more time to analyze key positions than do weaker players.

This thesis extends the work of Sigman *et al.* (2010) and includes a detailed examination of the time distributions for individual moves and the differences in the shapes of these distributions as a function of player skill. The purpose of this present thesis is to test an implication of key-position theory by examining the time distributions of stronger and weaker players (Study 1), and how the poor moves differ as a function of skill difference (Study 2).

As described previously, there is no clear definition of which positions are key-positions. However, if slow processes are required in only a few positions per game, one would expect there to be a very asymmetric distribution of times with some moves taking considerably more time than the majority of others. In contrast, recognition-action theory which assumes there is essentially no forward search when players are forced to move quickly, provides no clear reason that the distribution of move times under time pressure



should not be approximately symmetric. In other words, there is no reason to expect that some moves would take substantially longer than others.

As suggested by key-position theory, if stronger players do more calculation and spend more time on some key positions, one would expect that there to be an even more skewed distribution of move times for the stronger players than for the weaker players. In contrast, recognition-action theory, which assumes stronger players rely more on fast recognition-action processes to generate a move than weaker players, provides no reason for stronger players' distributions of move times to be more skewed. In other words, there is no reason to expect that the stronger players would take substantially more extreme moves and longer time on these moves than the weaker players.

In order to distinguish between these two possibilities and test the predictions from these two contrary views, Study 1 investigated the time distribution for individual moves during blitz chess and its functions of skill difference (5 minutes for the whole game for each player). Study 2 presented an analysis of poor moves as a function of skill difference.

### **Study 1: Move Times**

#### **Methods**

The subjects were selected from those playing publicly on the website of the Internet Chess Club (ICC; <http://www.chessclub.com>). This is a club with various strengths of players in different time control, including many of the best players in the world. The players receive a stable rating only after having played a sufficient number of games in a specific time control. Not only having over 200,000 active members, ICC also has invested a significant amount of effort into the detection of frauds: computer fraud

and having stronger players use others account. For the purpose of this study, the Elo ratings for the selected players must fall into these two ranges: the Elo range from 2300 to 2399 (stronger players) and the Elo range from 1600 to 1699 (weaker players). The selection of the players is based on the following protocol. First of all, the subjects, within Elo rating ranges, were selected if they were playing a blitz game during data collection days. Secondly, after a subject has been selected, we were looking into that player's play history, including this current game, and selecting the most recent game that the chosen player and the opponent both fall into the same Elo ranges. Based on this protocol, a total of 200 subjects (100 games) was selected. The stronger players in this present study have the Elo range from 2300 to 2397 ( $n = 100$ ;  $M = 2349$ ;  $SD = 31$ ); whereas the weaker players have Elo range from 1600 to 1699 ( $n = 100$ ;  $M = 1652$ ;  $SD = 28$ ).

#### *Time for Each Move*

In this present study, there are two ways of relative faster moves have been defined: "opening moves" and "time pressure moves." "Opening moves," are defined as moves 1-15 which often contain many "book moves," that are played from memory. In order to cover all the book moves, move 1 to 15 is a conservative estimation. "Time pressure moves" are defined as moves made when players had less than 30 seconds left. When the players do not have much time left, they must play fast. For the purpose of this study, I plotted and examined only the rest of moves, excluding "opening moves" and "time-pressure moves," which aimed to eliminate the relatively faster moves. Since these relatively faster moves have been excluded, one would expect that the time distribution presents a less skewed distribution. This present study is interested in not only the time

distribution for individual moves, but also the differences in the shapes of move time distributions as a function of player skill.

## Results and Discussion

The structure of the results and discussion section is as follows: First, I present the results of time distributions for moves. Next, I present comparisons of time distributions for players of different skill levels.

Figure 1 shows the time distribution for individual moves and its descriptive data from all subjects. As can be seen in this figure, the time distribution has an extremely positive skew. This asymmetrical time distribution shows there are some moves taking considerably more time than the majority of others. This time distribution is consistent with what would be expected if relatively deep search occurs only in some key positions and is not clearly predictable by recognition-action theory. No significance test was used to test for skewness because these scores are not independent. In order to run a significance test, the dimensionless third moment (skewness) was computed from each player in each game. The equation for this measure is:

$$Sk_M = \frac{n}{(n-1)(n-2)} \sum \left( \frac{X - Mean}{Standard\ Deviation} \right)^3$$

With each skewness score computed from one player in a game, these skewness scores can be considered sufficiently independent. Figure 2 shows the distribution of skewness across games. An examination of this figure shows that the positive skewness observed in Figure 1 is very unlikely to be due to sampling error since the vast majority of skewness scores are above 0. Further, a  $t$  test revealed that the mean skewness ( $M = 1.69$ ,  $SD = 0.91$ ) is significantly higher than 0,  $t(190) = 25.76$ ,  $p < .001$ . These three players who showed negative skewness were considered individually. First, these three players are all weaker

players. Secondly, these three players all made very few moves; two players made 5 moves and one player made 7 moves. In addition, one might argue that the extreme times may due to the subjects' distraction while playing the game; however, it unlikely occurs during the 5 minutes Blitz game.

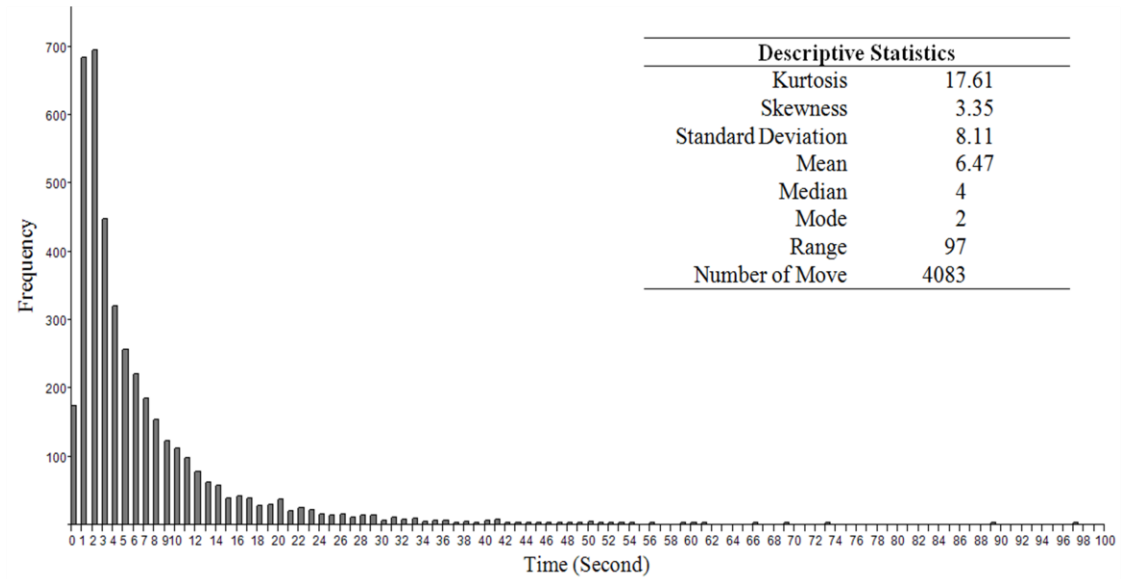


Figure 1. Time distribution of the moves from all subjects.

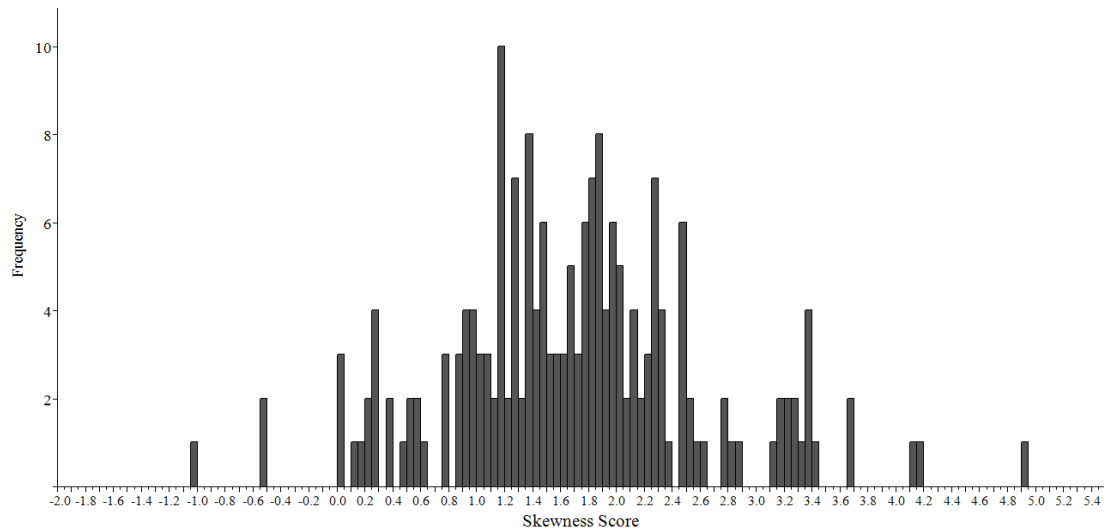


Figure 2. The frequency distribution of the skewness of the time distribution from each subject.

Recall that, according to recognition-action theory, there is essentially no forward search when players are forced to move quickly (Burns, 2004). This theory therefore provides no clear reason for the distribution of move times under time pressure to be highly skewed. In contrast, this extremely positive skewed distribution is just what one would expect from key-position theory.

Figure 3 shows the time distribution and its descriptive data of the moves from the stronger players and Figure 4 shows the time distribution and its descriptive data of the moves from the weaker players. Figure 5 shows the boxplots of time for each move in the two different skill level groups: stronger players versus weaker players. As can be seen in Figures 3, 4, and 5, the time distribution for the stronger players (mean skewness = 3.54) presents a more positive skewed distribution than the one for the weaker players (mean skewness = 3). The more skewed distribution of the stronger players shows that the stronger players took even more time in some positions than did the weaker players. The difference between these two time distributions is consistent with what would be expected by key-position theory: the stronger players spent more time in some key positions than do the weaker players even if they calculate fast. This finding may not be predictable by recognition-action theory.

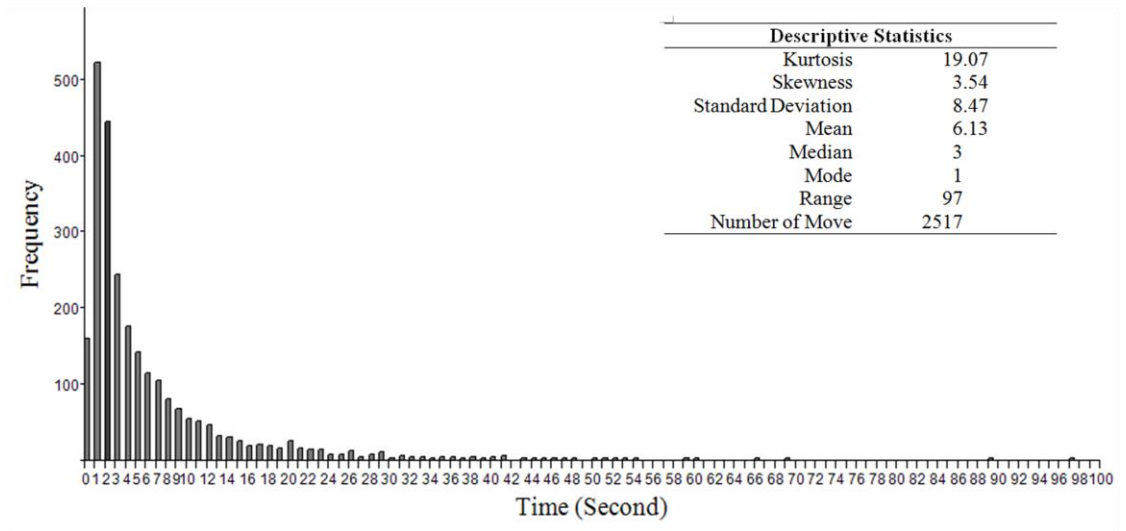


Figure 3. Time distribution of the moves from the stronger players.

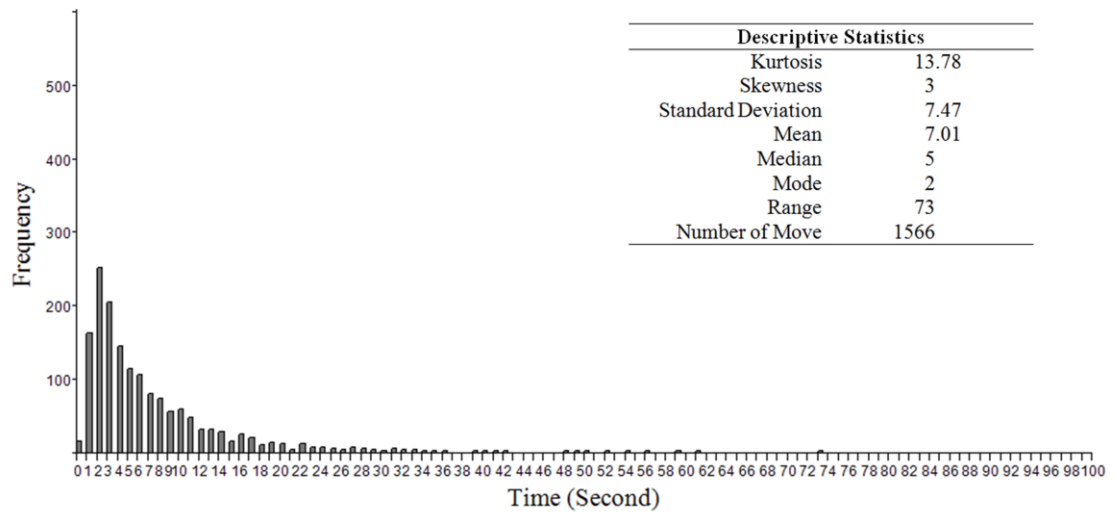


Figure 4. Time distribution of the moves from the weaker players.

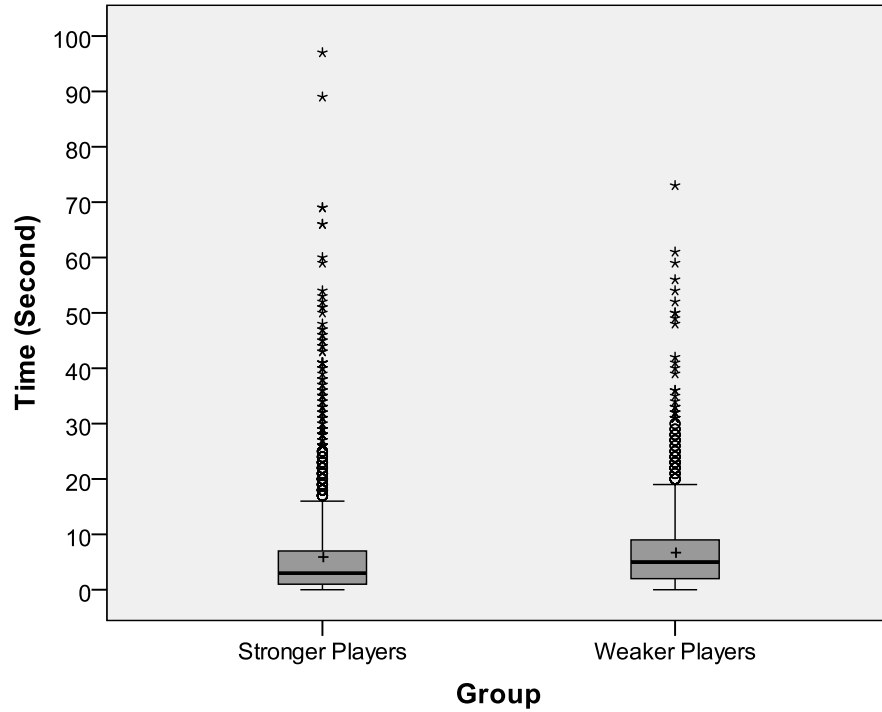


Figure 5. The boxplots of the two time distributions of the moves from the stronger players and the weaker players.

In order to run a significance test, the number of moves, skewness, standard deviation, mean, median, and maximum were computed from each player in each game. These data are presented in Table 1. One of most important results, consistent with key-position theory, is that the time distributions of the stronger players ( $M = 2.06$ ,  $SD = .81$ ) are more positively skewed than the distributions of the weaker players ( $M = 1.38$ ,  $SD = .86$ ),  $t(178) = 5.48$ ,  $p < .001$ . Secondly, the stronger players ( $M = 32.46$ ,  $SD = 16.03$ ) had significantly larger mean maximum than the weaker players ( $M = 23.93$ ,  $SD = 13.51$ ),  $t(178) = 3.85$ ,  $p < .001$ . A more detailed look at the maximum and skewness scores are shown in Figure 6 and Figure 7, which present the boxplots of the maximum and skewness scores as a function of playing strength. In addition, it is interesting to see that the stronger players ( $M = 26.88$ ,  $SD = 11.76$ ) made significantly more moves than the weaker players ( $M = 17.46$ ,  $SD = 8.62$ ),  $t(178) = 6.07$ ,  $p < .001$ . As shown in Table 1, the

stronger players made both more really fast moves and slow moves than the weaker players. Although there is no clear explanation for this observation, it is possible that the weaker players may reach the end of games sooner than the stronger players because they made game-ending blunders more often.

Table 1. The comparison of the time distributions of the moves for two different skill levels.

	<b>Stronger Players</b>		<b>Weaker Players</b>		<i>df</i>	<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Mean	6.92	3.10	7.17	2.76	178	-0.59	.558
Median	3.95	2.15	5.14	2.36	178	-3.56	<.001*
Maximum	32.46	16.03	23.93	13.51	178	3.85	<.001*
Standard Deviation	8.23	4.48	6.50	3.69	178	2.83	.005*
Skewness	2.06	.81	1.38	.86	178	5.48	<.001*
Number of Moves	26.88	11.76	17.49	8.62	178	6.07	<.001*
Proportion of Outside Values	.042	.04	.028	.04	178	2.28	.024*
Proportion of Far Out Values	.029	.04	.014	.03	178	2.98	.003*

Note: There were 87 weaker players and 93 stronger players have been analyzed in this table. We have excluded the subjects with any one of the missing data among the descriptive data, such as mean, median, mode, maximum, standard deviation, and skewness. The missing data is because the player has generated less than 5 moves.



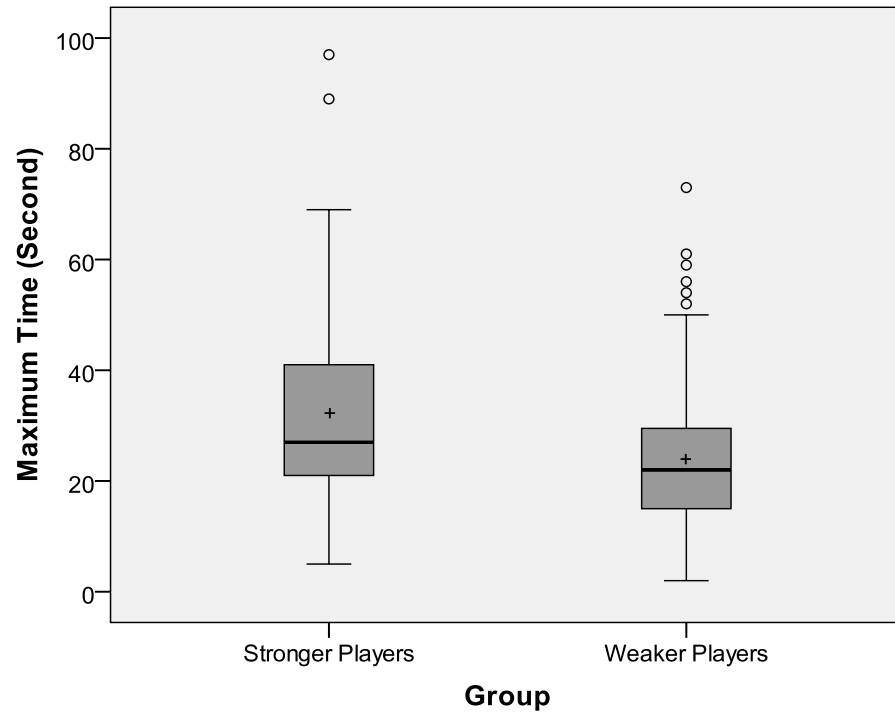


Figure 6. The boxplots of the maximum of the time moves from the stronger players and the weaker players.

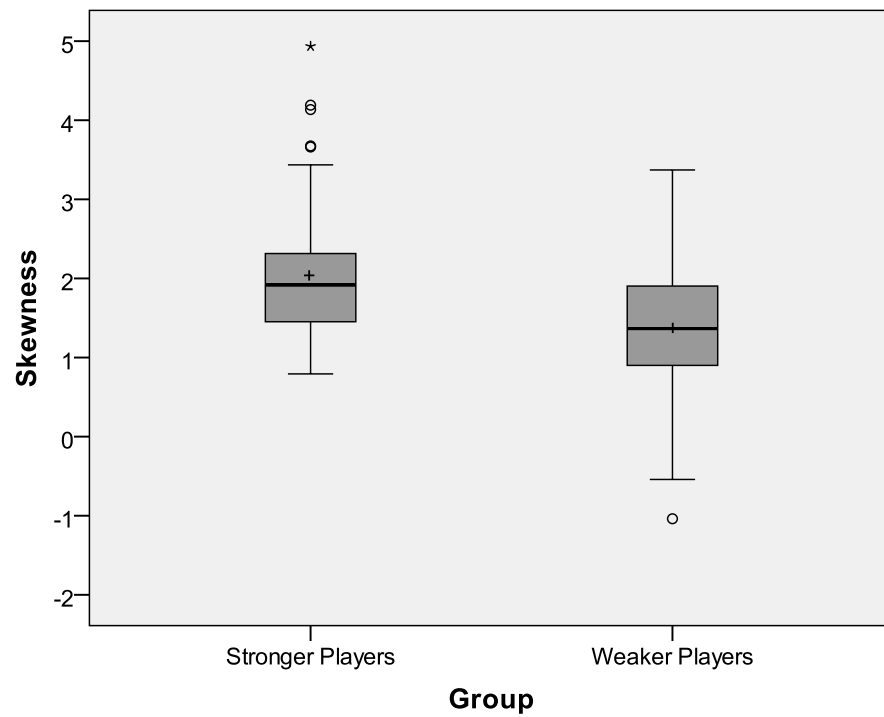


Figure 7. The boxplots of the skewness of the time distributions from the stronger players and the weaker players.

I also identified and counted the number of the outside values (the value from 1.5 IQR to 3.0 IQR above the 75<sup>th</sup> percentile) and far out values (above 3.0 IQR above the 75<sup>th</sup> percentile) of the time spent on each move based on individual subject. I compared the proportion of outside values and far out values for the two different skill levels. For each player, the proportion of his or her moves that were outside and far out values were computed. The strong players had a significantly higher mean proportion of outside ( $M = .042$ ,  $SD = .04$ ) and far out values ( $M = .029$ ,  $SD = .04$ ) than the weaker players (Outside:  $M = .028$ ,  $SD = .04$ ; Far Out:  $M = .014$ ,  $SD = .0$ ),  $F(1, 178) = 14.58$ ,  $p < .001$ . One factor affecting the proportion of outside and far out values is the number of relatively fast moves played. The fact that the stronger players made somewhat more fast moves complicates the interpretation of these results. However, the finding that the stronger players had significantly larger maximum times than the weaker players is not subject to this problem.

In sum, recognition-action theory provides no reason to expect that the stronger players make more slow moves and take longer time on the slow moves than the weaker players. In contrast, key-position theory posits that skill differences in depth of search depend on the positions, and the stronger players are able to do more calculations and they may spend more time in key positions even if they calculate faster. According to this theory, one would expect that the stronger players spend longer time on some positions even in speed chess when time is severely limited than the weaker players. Therefore, the finding that the time distribution for the stronger players presents an even more positively skewed distributions than the time distribution for the for weaker players indicates that the stronger players take considerably more time than the weaker players in some

positions, presumably those for which the deep search is required. The findings that stronger players made more slow moves and took longer on the slowest moves than the weaker players is consistent with key-position theory.

## **Study 2: Skill Differences in Poor Moves**

It is generally very difficult to assess how far a player looked ahead on a given move from only the game score. However, the analysis of poor moves provides an indirect way to do this. By analyzing the position in which each poor move was made, one can determine how many plies the player would have had to look ahead correctly in order to have seen that it was an error. Therefore, the number of plies in poor moves can be interpreted as an indirect measure of depth of search. As applied to skill differences, if stronger players search deeper than weaker players, then their poor moves will tend to be higher-ply poor moves than those poor moves that made by the weaker players. If the weaker players and the stronger players are looking ahead in the same degree, there is no reason to think that the poor moves made by the stronger players have higher plies than the poor moves made by the weaker players.

In this study, differences in these poor moves as function of player skill levels were analyzed. I classified the poor moves in terms of how many plies the player would have had to look ahead correctly in order to have avoided it.

## **Methods**

### *Poor Moves*

In order to identify all the major poor moves made, I analyzed the games by means of the game-analysis facility of the Fritz 13 chess playing program (ChessBase, Hambury, Germany). This program has long been one of the world's strongest commercially available computer chess programs. Three criteria were used to identify

and to define the poor moves. The criteria for identifying the poor moves are very similar with the ones that Chabris and Hearst (2003) applied to define the “true blunder” in their study. I called these identified moves as “poor moves” instead of blunders in this study, because they are not stupid, ignorant, or careless mistakes (Merriam-Webster Online Dictionary). It is possible that these poor moves could have been made just because these positions were complex and difficult. If one move requires the player to look 10-ply ahead in that position, this move is just a poor move rather a blunder. However, based on the definition of blunders, one could call these 2-ply and 1-ply poor moves as “blunders.”

The first two criteria are the same with the ones that Chabris and Hearst (2003) used in their study. First, by processing each move in the Fritz 13, I was able to obtain an objective evaluation of each move. By setting up the threshold as 10 ply, this program analyzed every move with a nominal 10-ply exhaustive search, and this depth of search starts at the opponent’s side. It means that this program looked five moves ahead by each player in each position, and this program chose a best move. If the actual move played is evaluated as at least 1.5 pawns worse than the program’s choice for the best move, this move was defined as the “candidate poor move.” The 1.5-pawn criterion is not chosen arbitrarily, since this size is generally considered as a presumed threshold that a game is theoretically to be sufficient to win (see, e.g., Hartmann, 1989). In addition, the 10-ply threshold was chosen not only to be consistent with Chabris and Hearst’s (2003) study, but also because it is unlikely that the players will search deeply beyond the 10 plies, especially in speed chess. Secondly, if the same side retained a prior advantage of at least 3.0 pawns in the position scores even after a candidate poor move, this move has been excluded. Besides these two criteria, I also have added a new criterion which is that if the

opposite side retained an advantage in the position scores at least 3.0 pawns, and then the candidate poor move has also been excluded. The logic is that if a player still has a clearly winning position or losing position after a poor move, this poor move has no practical consequences. After applying these three criteria, the remaining moves were all considered as “poor moves.”

In addition, I have performed a more elaborate analysis on these defined poor moves. To be more specific, if one move has been identified as a poor move, it means that this move was at least 1.5 pawns worse than the program’s choice for the best move with 10 plies search. Then, I continually process this move with 9 ply threshold, which means that the program will choose a best move by looking 9 plies (4.5 moves) ahead. If this actual move played is evaluated as at least 1.5 pawns worse than the program’s choice for the best move, then, I continually process it with 8 ply threshold. I kept processing this move with one degree decreasing threshold until the program no longer identified this move as a poor move. If the program does not identify the actual move played as a poor move when the threshold is set up in 5 ply, then I call this move as a 6-ply poor move. It means that if the player would have had to look 6 plies ahead, then this poor move could have been avoided. The lowest threshold I applied was 1 ply. If a move has been identified as 2-ply or 1-ply poor move, then I called it as blunder. Therefore, by setting up different thresholds (e.g., 9 ply) in the Fritz 13, I was able to classify poor moves in terms of how many plies the player would have had to look ahead correctly in order to have avoided it.

In order to clarify the definition of n-ply poor move, Figure 8 and 9 illustrate the examples of how 1-ply and 2-ply poor move have been identified and what these mean.

In the example of Figure 8, based on the criteria and default 10-ply threshold, Black's "Rxf1" move (as indicated by the black arrow) was initially been identified as a poor move. The program chose the move "Qxf1" as the best move instead. Since this move, "Rxf1," has been identified as a poor move, I kept processing it with one degree decreasing threshold until the program no longer identified it. In this particular move, the program has identified it as a poor move from 10-ply to 1-ply threshold. Therefore, I called it as 1-ply poor move. To be specific, it means that if Black would have looked 1-ply beyond Rxf1, White's "Qe8 checkmate" move (as indicated by the red arrow), then Black could have avoided being checkmated in the next ply. In the example of Figure 9, White's "Nc3" move (as indicated by the black arrow) was initially identified as a poor move. By applying the similar procedures described above, this move was classified as 2-ply poor move. If White would have looked 2-ply ahead, Black's "Bd2" and the following White's "Ne2" (as indicated by the red arrow), then White's rook could avoid being captured in the next ply by "Bxe1."



Figure 8. An example for Black's 1-ply poor move.

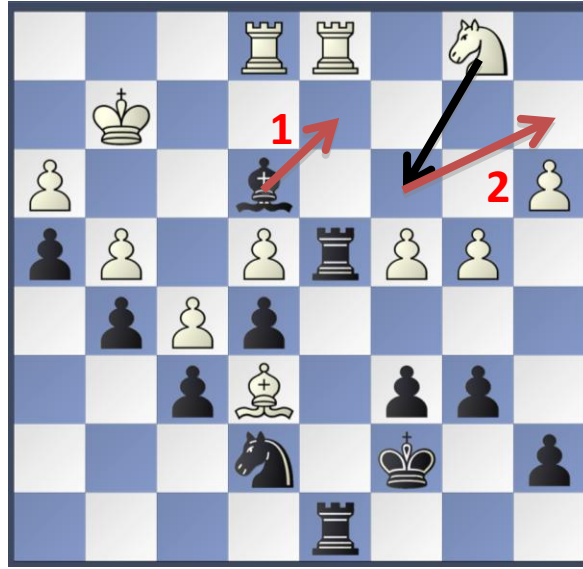


Figure 9. An example for White's 2-ply poor move.

## Results and Discussion

The first analysis on these poor moves compared the proportion of poor moves for the two different skill levels based on individual players. The proportion of poor moves is the number of the poor moves divided by the total number of the moves had been made by each individual player. As expected, the stronger players ( $M = .048$ ,  $SD = .064$ ) made proportionally fewer poor moves than the weaker players ( $M = .092$ ,  $SD = .129$ ),  $t(194) = -2.994$ ,  $p = .003$ .

Next, I compared the ply level of the poor moves for two different skill levels. As described previously in the methods section, each poor move was identified as an n-ply poor move, which in terms of the player would have had to look ahead n plies to avoid that poor move. However, conducting a significance test on the data is problematical because some players made more than one poor move and therefore these data are not independent. In order to run a significance test, I randomly selected one poor move if a

player made more than one poor move in a game. With each n-ply poor move obtained from one player in a game, the errors can be considered sufficiently independent. Figure 10 shows a back-to-back dot-plot of the randomly selected one n-ply poor moves from each player for the stronger and the weaker players. As expected, these two distributions differ. Table 2 shows the proportion of n-ply poor moves of total number of the moves had been made for two different skill levels. The stronger players (proportion = .009) made proportionally fewer blunders (2-ply or 1-ply poor moves) than the weaker players (proportion = .024). In addition, even though the stronger players made many poor moves, their poor moves tended to be higher-ply poor moves (i.e., 9-ply and 10-ply). Since these two distributions are highly skewed as shown in Figure 10, I applied the log transformation to the number of plies. After the log transformation, the weaker players (*Geometric Mean* = 2.67) made significantly lower-ply poor moves than the stronger players (*Geometric Mean* = 3.74),  $t(122) = 3.76, p < .001$  (95% CI on the Arithmetic mean difference: 1.18 to 1.69). This suggests that even when stronger players made a poor move, they searched deeper than weaker players. Defining a blunder as a 2-ply or 1-ply poor move, the stronger players made 23 blunders and 38 higher-ply poor moves, whereas the weaker players made 38 blunders and 25 higher-ply poor moves. The proportions of poor moves were significantly different,  $\chi^2(1, N = 124) = 6.34, p = .012$ . This supports the view that the stronger players search deeper than the weaker players.



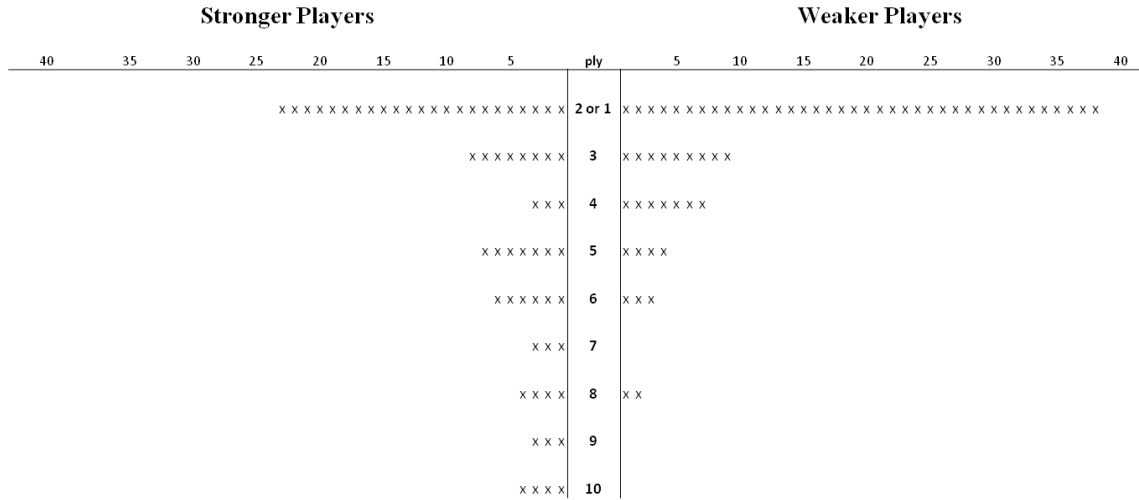


Figure 10. The back-to-back dot plot of the n-ply poor move for the stronger and the weaker players.

Table 2. The proportion of n-ply poor moves over all moves for two different skill levels.

n-ply poor move	Stronger Players		Weaker Players	
	<i>N</i>	<i>Proportion</i>	<i>N</i>	<i>Proportion</i>
2- or 1- ply	23	0.0091	38	0.0243
3- ply	8	0.0032	9	0.0057
4- ply	3	0.0012	7	0.0045
5- ply	7	0.0028	4	0.0026
6- ply	6	0.0024	3	0.0019
7- ply	3	0.0012	0	0
8- ply	4	0.0016	2	0.0013
9- ply	3	0.0012	0	0
10- ply	4	0.0016	0	0

Note: The stronger players made a total of 2,517 moves, whereas the weaker players made totally 1,566 moves.

In sum, the stronger players made proportionally fewer low-ply poor moves and proportionally more high-ply poor move than the weaker players. Although even the

stronger players made poor moves, these moves tended to be higher-ply than those made by the weaker players. These findings are consistent with the view of key-position theory that the stronger players calculate more and search deeper than the weaker players even under time pressure.

### **General Discussion**

De Groot's (1946) finding of no difference in the depth of search between stronger and relatively weaker players has been frequently cited in studies of expertise and highlighted in many cognitive psychology textbooks. Many studies have replicated de Groot's study and interpreted the findings as evidence supporting recognition-action theory. However, it is surprising to see most of studies replicated de Groot's study by using the same position, and, of course, it led to the same finding: no difference in the depth of search between stronger and relative weaker players. Based on these findings, some studies infer that pattern recognition play a more important role than the depth of search in chess expertise. For example, "the fact that the differences in search depth between master and intermediate players...were not statistically significant and the effect sizes were small suggests that other factors, such as pattern recognition, continue to play a more important role than search depth in chess expertise" (Connors *et al.*, 2011, p. 9). However, this frequently cited and accepted conclusion is made by only few positions, especially the position A used in de Groot's study. As shown in the work of Bilalic *et al.* (2009) and the statement of Kasparov (2010), the skill difference in the depth of search must depend on the positions. Therefore, I would argue that studies concluding that there are no skill differences in depth of search have generalized their findings from the few positions investigated beyond what is warranted.

In this thesis, I have proposed “key-position theory,” which states that the slow processes, such as searching deeply and planning, are required only in a few key positions, which occur even when the time for calculation is severely limited. The highly-positive skewed distributions reported in Study 1 are consistent with the predictions of key-position theory, and these distributions are difficult to be accounted for recognition-action theory. Moreover, the stronger players spent more time on some positions than did the weaker players. This implies that these players were able to do extensive calculation and searching even when time is severely limited. In addition, Study 2 provides indirect evidence that the stronger players are able to search deeper and calculate more than the weaker players. Specifically, they made proportionally few blunders (2-ply poor moves) than the weaker players. Although the stronger players made poor moves, these poor moves tended to be higher-ply poor moves. In sum, both studies provide explicit and implicit evidence for the fact that there are some slow processes involved in speed chess play. Despite its problems accommodating the present findings, recognition-action theory is still important because it provides a way to generate candidate moves. However, it appears that the improvement with experience in recognition-action mechanism is necessary but not sufficient for players to reach a high level of expertise. It would not be prudent to ignore the reflections of great chess players and teachers on the factors involved in chess skill, such as searching deeply, planning, and knowing principles. Regarding depth of search, Kotov (1987) stated that “having examined the games of other players, particularly masters.... I became even more convinced that the ability to analyse clearly a sufficient number of variations so as to clarify the position was the basic condition for success” (p. 16). Regarding planning, Kotov (1987) stated that “one factor

is always present in all a grandmaster does. He always takes account of it when planning for the immediate or the distant future” (p. 147). Regarding knowing the principles of positional play, Nimzowitsch’s (1925) book “My System” introduced several very influential rules and principles of chess strategy, such as the outpost, liquidation to avoid embarrassment, and mysterious rook moves. This book is generally considered to be one of the most important chess books.

Therefore, these data provide evidence challenging the view that chess expertise is solely a function of practice (Ericsson, 2006, Simon & Chase, 1973). These authors argued that a player can become a master simply by engaging in thousands of hours of directed practice. According to them, such practice allows the players to build up the patterns of chess positions required for recognition-action mechanism. The current findings suggest that the widely-accepted explanation of the development of chess expertise in terms of the building up of thousands of patterns of chess pieces is not sufficient. Therefore, in addition to being familiar with many patterns of chess pieces, it is likely that factors such as the ability to search deeply, plan strategically, and apply positional principles are critical determinants of players’ skill levels. In sum, in agreement with the conclusion of Gobet (2008) the role of practice for achieving the high levels of expertise is necessary but not sufficient.

## References

- Bilalić, M., McLeod, P., & Gobet, F. (2008). Expert and 'novice' problem solving strategies in chess: Sixty years of citing de Groot (1946). *Thinking and Reasoning*, 14(4), 395-408.
- Bilalić, M., McLeod, P., & Gobet, F. (2009). Specialization effect and its influence on memory and problem solving in expert chess players. *Cognitive Science*, 33(6), 1117–1143.
- Blunder. 2012. In *Merriam-Webster.com*. Retrieved October 14, 2012, from <http://www.merriam-webster.com/dictionary/blunder>.
- Burns, B. D. (2004). The effects of speed on skilled chess performance. *Psychological Science*, 15, 442–447.
- Calderwood, R., Klein, G. A., & Crandall, B. W. (1988). Time pressure, skill, and move quality in chess. *American Journal of Psychology*, 101, 481–493.
- Campitelli, G., & Gobet, F. (2004). Adaptive expert decision making: Skilled chess players search more and deeper. *International Computer Games Association Journal*, 27, 209–216.
- Chabris, C. F., & Hearst, E. S. (2003). Visualization, pattern recognition, and forward search: Effects of playing speed and sight of the position on grandmaster chess errors. *Cognitive Science*, 27, 637–648.
- Chabris, C. F., & Hearst, E. S. (2005). Search, recognition, and visualization in chess: Rebuttal to Gobet's critique of Chabris & Hearst (2003). Unpublished manuscript.
- Charness, N. (1976). Memory for chess positions: Resistance to interference. *Journal of Experimental Psychology: Learning and Memory*, 2, 641-653

- Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 467–476.
- Charness, N. (1989). Expertise in chess and bridge. In D. Klahr, & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 183–208). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Charness, N. (1991). Expertise in chess: The balance between knowledge and search. In K. A. Ericsson, & J. Smith (Eds.), *Towards a general theory of expertise: Prospects and limits* (pp. 39-63). England: Cambridge University Press.
- Chase, W. G., & Simon, H. A. (1973a). Perception in chess. *Cognitive Psychology*, 4, 55–81.
- Chase, W. G., & Simon, H. A. (1973b). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.
- Connors, M. H., Burns, B. D., & Campitelli, G. (2011). Expertise in complex decision making: The role of search in chess 70 years after de Groot. *Cognitive Science*, 35(8), 1567-1579.
- Cooke, N. J., Atlas, R. S., Lane, D. M., & Berger, R. G. (1993). Role of high-level knowledge in memory for chess positions. *American Journal of Psychology*, 106(3), 321-351.
- Craik, F. I. M., & Lockhart, R. S. (1972). Level of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- de Groot, A. D. (1946). *Het denken van den schaker*. Amsterdam: Noord Hollandsche.
- de Groot, A. D. (1965). *Thought and choice in chess*. The Hague: Mouton Publishers.
- de Groot, A. D. (1978). *Thought and choice in chess* (2nd ed.). The Hague: Mouton

Publishers.

- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J., Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683 – 703). Cambridge, England: Cambridge University Press.
- Frey, P. W., & Adelman, P. (1976). Recall memory for visually presented chess position. *Memory & Cognition*, 4(5), 541-547.
- Gobet, F. (1997). A pattern-recognition theory of search in expert problem solving. *Thinking and Reasoning*, 3(4), 291-313.
- Gobet, F. (1998). Chess players' thinking revisited. *Swiss Journal of Psychology*, 57, 18-32.
- Gobet, F. (2003). Forward search, pattern recognition and visualization in expert chess: A reply to Chabris and Hearst (2003). Unpublished manuscript.
- Gobet, F. (2008). The role of deliberate practice in expertise: Necessary but not sufficient. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (p. 823). Mahwah, NJ: Erlbaum.
- Gobet, F., & Simon, H. A. (1996a). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, 31, 1-40.
- Gobet, F., & Simon, H. A. (1996b). The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grand-master-level chess. *Psychological Science*, 7(1), 52-55.
- Gobet, F., & Simon, H. A. (2000). Reply to Lassiter. *Psychological Science*, 11, 174.

- Gruber, H. (1991). *Qualitative aspekte von expertise im schach. [Qualitative aspects of expertise in chess]*. Aachen: Feenschach.
- Hartmann, D. (1989). Notions of evaluation functions tested against grandmaster games. In D. F. Beal (Ed.), *Advances in computer chess 5* (pp. 91-141). Amsterdam: North-Holland.
- Holding, D. H. (1979). The evaluation of chess positions. *Simulation and Games*, 10, 207-221.
- Holding, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Erlbaum.
- Holding, D. H. (1992). Theories of chess skill. *Psychological Research*, 54, 10-16.
- Holding, D. H., & Pfau, H. D. (1985). Thinking ahead in chess. *American Journal of Psychology*, 98, 271-282.
- Holding, D. H., & Reynolds, R. I. (1982). Recall or evaluation of chess positions as determinants of chess skill. *Memory and Cognition*, 10, 237-242.
- Kasparov, G. (1993a, September). Simuls in my life. *Chess Life*, 732-733.
- Kasparov, G. (1993b, October). Simuls in my life. *Chess Life*, 802-803.
- Kasparov, G. (1993c, November). Simuls in my life. *Chess Life*, 890-891.
- Kasparov, G. (1993d, December). Simuls in my life. *Chess Life*, 972-973.
- Kasparov, G. (1994, January). Simuls in my life. *Chess Life*, 6-8.
- Kasparov, G. (2010, February 11). The chess master and the computer. *The New York Review of Book*, 205.
- Kotov, A. (1987). *Think like a grandmaster*. London: Batsford.
- Lassiter, G. D. (2000). The relative contributions of recognition and search-evaluation



- processes to high-level chess performance: Comment on Gobet and Simon.  
*Psychological Science*, 11, 172–173.
- Lane, D. M., & Robertson, L. (1979). The generality of the levels of processing hypothesis: An application to memory for chess position. *Memory & Cognition*, 7(4), 253-256.
- Moxley, J. H., Ericsson, K. A., Charness, N., & Krampe, R. T. (2012). The role of intuition and deliberative thinking in experts' superior tactical decision-making. *Cognition*, 124(1), 72-78.
- Nimzowitsch, A. (1925). *Mein system*. Berlin: Verlag B. Kagan.
- Reshevsky, S. (1945). *Reshevsky on chess: The U.S. champion tells how he wins*. NY: Chess Review.
- Reshevsky, S. (1948). *Reshevsky's best games of chess*. NY: Dover Publication Inc.
- Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading* (pp. 265-303). Hillsdale, NJ: Erlbaum.
- Saariluoma, P. (1990). Apperception and restructuring in chess player's problem solving. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie & G. Erdos (Eds.), *Lines of thinking: Reflection in the psychology of thinking. Vol II*. (pp. 41-57). London: Wiley.
- Saariluoma, P. (1992). Error in chess: The apperception-restructuring view. *Psychological Research*, 54, 17-26.
- Sigman, M., Etchemendy, P., Slezak, D. F., & Cecchi, G. A. (2010). Response time distributions in rapid chess: A large-scale decision making experiment. *Frontiers*

*in Neuroscience, 4, 1-12.*

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 6*(4), 394-403.

Simon, H. A., & Gilmarin, K. (1973). A simulation of memory for chess positions.

*Cognitive Psychology, 5, 29-46.*

Vicente, K. J., & de Groot, A. D. (1990). The memory recall paradigm: Straightening out the historical record. *American Psychologist, 45*(2), 285-287.

van Harreveld, F., Wagenmakers, E.-J., & van der Maas, H. L. J. (2007). The effects of time pressure on chess skill: An investigation into fast and slow processes underlying expert performance. *Psychological Research, 71, 591-597.*